AD-A046 878 NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN D--ETC F/G 5/3
FORECASTING THE NAVAL ENLISTED PERSONNEL FORCE STRUCTURE TO EST--ETC(U)
NOV 77 M CHIPMAN
UNCLASSIFIED NPRDC-TR-78-4 NL

1 OF 1
ADA046878

END
DATE
FILMED
12-77
DDC

AD A046878

NPRDC TR 78-4

November 1977

FORECASTING THE NAVAL ENLISTED PERSONNEL FORCE STRUCTURE TO ESTIMATE BASIC PAY

Mark Chipman

Reviewed by
Joe Silverman

Approved by
James J. Regan
Technical Director

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) NPRDC-TR-78-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (2) FORECASTING THE NAVAL ENLISTED PERSONNEL FORCE STRUCTURE TO ESTIMATE BASIC PAY.		5. TYPE OF REPORT & PERIOD COVERED Final rept. Feb [redacted] - Oct [redacted]77,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Mark Chipman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152 (Code 303)		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 73C201 ZF55-521-010-03.11
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152 (Code 303)		12. REPORT DATE (11) Nove[redacted]77
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (12) 32p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(16) F55521

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(17) ZF55521010

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Forecasting
Force Structure
Time Series Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A primary concern of Navy management is the ability to meet the Navy's manpower requirements in terms of both the quantity and quality of personnel. Recruitment, promotion, retirement, and other managerial policies are all directly related to and restricted by the Military Personnel, Navy (MPN) budget. Because over 50 percent of the MPN budget is allocated for enlisted basic pay, special attention must be given to the accurate forecasting of enlisted populations. The Naval Pay Predictor, Enlisted (NAPPE) Model,

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390772

designed to accomplish this objective, relies solely upon historical population data and user-supplied enlisted end-strengths. Time series analysis is used to determine a general set of forecasting models that adequately explain the historical data. Other statistical procedures, including those employed in costing the enlisted force and in estimating recruit input populations, are also detailed. Validation results indicating errors of less than .1 percent for total enlisted basic pay are presented.

FOREWORD

The effort described in this report supports the development of management decision models, an exploratory development objective under Task Area ZF55.521.010. The overall objective of this task area is to develop techniques to improve the Navy's managerial decision-making capabilities. The main effort in FY77 is directed toward solving two problems. The first concerns the flow of recruits through the lower three pay grades (E-1 through E-3) into the petty officer pay grades (E-4 through E-9). The objective is to assure sufficient input to the petty officer force structure so as to minimize future personnel shortages and surpluses. The second problem is the forecasting of the Navy's military manpower budget to avoid cost overruns and personnel turbulence. This problem has been addressed by developing the Naval Personnel Pay Predictor, Enlisted (NAPPE) model, which is currently being used by the Bureau of Naval Personnel to forecast enlisted basic pay and to distribute the enlisted force over 31 length-of-service categories.

Acknowledgments are due to Joe Silverman, Dr. Kenneth Leland, and Melvyn Moy of the Navy Personnel Research and Development Center, San Diego; Duane Kirkland of IBM, Rochester, Minnesota; and the staff of the Active Enlisted Plans Division in the Bureau of Naval Personnel, Washington, D.C.

J. J. CLARKIN
Commanding Officer

SUMMARY

Problem

Navy planning and policy decisions are marked by complex problems and multiple, conflicting objectives. One such problem is that of forecasting the Navy's military manpower budget to avoid cost overruns and personnel turbulence. In particular, cost overruns can occur due to inaccurate estimates of either the personnel force structure (personnel arrayed by length-of-service (LOS) and pay grade) or the various pay rates. To control for the first source of error, forecasting models are being developed to predict required obligations based on a desired or allocated man-year average by pay grade.

Objective

The objective of this effort was to provide a detailed description of the analysis and procedures used to formulate a model for forecasting Navy enlisted force structures and basic pay obligations. The model, which is being used to monitor the age characteristics of the force and its costs in basic pay, is formally known as the Naval Personnel Pay Predictor, Enlisted (NAPPE).

Approach

NAPPE relies upon historical USN, USNR, and All Navy (ALNAV) quarterly force structure files dating back to 1957. Time series analysis techniques were applied to these files to find a particular set of time series models that would be appropriate for forecasting the LOS marginal distribution of each array. Various combinations of these forecasts were then compared to obtain a "best" forecast for the ALNAV LOS distribution. Additional statistical procedures were developed for (1) deriving the interior of the force structure matrix given the forecasted LOS and inputted pay grade marginal distributions, (2) forecasting the force structure for personnel with less than 1 year of service or more than 30 years of service, (3) costing the force structure, (4) estimating average strength, and (5) validating the model.

Findings

The statistical techniques employed in NAPPE proved to be highly reliable in producing estimates of enlisted basic pay. Validation results indicated forecasting errors of less than .2 percent for FY75, FY76, the quarter ending in September 1976, and the first 3 quarters of FY77. NAPPE's predictions for mean LOS of the force also indicated a high degree of accuracy. Generally, though, the accuracy of the forecasts diminished as the forecast lead time increased.

Conclusions

1. Sufficient data are available to allow time series analysis techniques to be applied toward the development of a model to forecast the enlisted force structure, and hence, basic pay obligations.

2. A particular set of time series models exists that adequately explains a great majority of the data. This set of models provides a solid statistical basis for making force structure forecasts.

3. The NAPPE model can be utilized by Navy management to provide future basic pay and mean LOS estimates.

Recommendations

1. Future work on NAPPE should concentrate on developing an interactive computer program, making it more responsive to user needs.

2. NAPPE should be expanded to include forecasts of other enlisted budgetary components, such as basic allowance for quarters, proficiency pay, clothing allowances, etc.

3. The Navy should develop a Naval Pay Predictor, Officer (NAPPO) model to forecast officer basic pay and, eventually, other officer costs.

CONTENTS

TABLES

FIGURES

INTRODUCTION

Problem

The Navy's ability to meet its manpower requirements in terms of both quantity and quality of personnel, a primary concern of Navy management, is restricted by the size of the personnel budget. Recruitment, promotion, retirement, and other managerial policies are all directly or indirectly related to the amount of dollars available in the Military Personnel, Navy (MPN) budget, over half of which is allocated for basic pay of enlisted personnel. In fiscal year 1976, $2,942,985,000 of obligations were incurred for this one item. Thus, even a 1 percent discrepancy in forecasting the enlisted basic pay component may result in an over- or underexpenditure of $30 million, an amount that could wreak havoc with the Navy's personnel policies. Because the amount of basic pay is directly proportional to the enlisted force structure (i.e., the number of personnel tabulated by pay grade and length of service), the force structure's configuration must be forecasted as accurately as possible in order to closely predict future basic pay. A previous report (Chipman & Wilson, Note 1) describes the current ability of Navy management to project the cost of enlisted basic pay and the other 10 categories of Budget Activity Two.

Background

The enlisted force structure is characterized by 9 pay grades and 31 length-of-service (LOS) categories. A representative force structure matrix is depicted in Figure 1. The pay grades are composed of the nonrated force (E-1 through E-3) and the six petty officer ratings (E-4 through E-9). The LOS category refers to the number of years in service, with cell 31 including all personnel with more than 30 years of service. Cells 1 through 4 generally contain first-term personnel, and cells 5 through 31, personnel who are often referred to as careerists.

Once a force structure matrix is projected, a simple procedure is followed to obtain the cost associated with that configuration. First, because the force structure fluctuates over time, the projected population for a certain period of time is obtained by finding the average of the numbers of personnel within a pay category at the beginning and at the end of the period. Next, the cost is calculated by multiplying the average strength within each pay category by a prespecified statutory rate and summing over all pay categories. The 62 pay categories and a set of representative rates are depicted in Figure 2. It is interesting to note that, for pay grades E-1 and E-2, cumulative years of service have no effect upon the pay scale. Thus, E-1 itself is one pay category, as is E-2. The primary task, then, in forecasting the enlisted basic pay entails the forecasting of the force structure.

To allow Navy managers to improve their predictions, the Naval Personnel Pay Predictor, Enlisted (NAPPE) model was designed by the Navy Personnel Research and Development Center for use in forecasting the enlisted force structure and the corresponding budget required to maintain various levels of future authorized strength.

All Navy Inventory for 6/74

LOS	E1	E2	E3	E4	E5	E6	E7	E8	E9	TOTAL
1	25183	33026	8758	1985	28	1	0	0	0	71981
2	4734	30367	35352	13828	144	5	1	0	0	84431
3	1145	7619	22430	29275	3644	6	1	0	0	64120
4	186	1325	7807	25134	12597	47	0	0	0	47096
5	59	241	2379	9688	12565	421	1	0	0	25354
6	19	88	1612	3817	12182	1438	4	0	0	19160
7	26	78	2463	2178	8360	2427	7	0	0	15539
8	7	24	770	1963	5243	2307	12	0	0	10326
9	3	12	316	1175	5314	3954	42	0	0	10816
10	1	9	147	960	4532	3858	130	0	0	9637
11	1	3	82	522	3222	4344	459	0	0	8633
12	0	2	40	258	2237	4426	773	5	0	7741
13	1	0	19	217	2063	5262	1208	53	0	8823
14	1	1	35	327	1959	5637	1798	134	5	9897
15	1	2	13	187	1403	5195	2313	310	27	9451
16	0	1	12	131	1038	4785	2905	532	63	9467
17	0	1	6	81	786	4438	3416	727	130	9585
18	1	0	4	55	570	4035	3520	893	200	9278
19	0	0	7	121	1165	5642	5157	1288	335	13715
20	0	2	5	94	892	4316	4800	1372	481	11962
21	0	0	2	26	284	1753	2296	804	295	5460
22	0	0	0	12	177	928	1471	600	276	3464
23	0	0	0	5	88	594	1189	572	381	2829
24	0	0	2	0	26	167	478	274	203	1150
25	0	0	1	2	24	97	234	113	100	571
26	0	0	1	1	6	58	282	190	183	721
27	0	0	1	0	4	58	314	205	240	822
28	0	0	1	1	3	36	178	136	153	508
29	0	0	0	0	7	50	276	191	224	748
30	0	0	0	0	8	49	203	171	219	650
31	0	1	0	1	6	36	272	230	253	799
TOTAL	34368	72802	82265	92044	80577	66370	33740	8800	3768	474734

Figure 1. A typical force-structure matrix showing population by pay grade and longevity.

Pay Grade	Cumulative Years of Service													
	Under 2	Over 2	Over 3	Over 4	Over 6	Over 8	Over 10	Over 12	Over 14	Over 16	Over 18	Over 20	Over 22	Over 26
E-9						←	1055.40	1079.40	1104.00	1129.50	1154.10	1176.90	1239.00	1359.00
E-8					←	885.60	910.20	934.50	959.10	984.00	1006.80	1031.70	1092.00	1214.10
E-7	618.30	667.20	692.10	716.10	741.00	764.10	788.40	813.30	849.90	873.90	898.50	910.20	971.40	1092.00
E-6	534.00	582.30	606.60	631.80	655.50	679.80	704.40	741.00	764.10	788.40	800.70	→		
E-5	468.90	510.30	534.90	558.30	594.60	618.90	634.80	667.20	679.80	→				
E-4	450.60	475.80	503.70	543.00	564.30	→								
E-3	433.20	457.20	475.50	494.40	→									
E-2	417.30	→												
E-1	374.40	→												

Figure 2. A typical monthly pay matrix for enlisted personnel.

The NAPPE model employs several sources of information. First, three separate quarterly force structure files dating back to 1957 are maintained. The three files contain USN, USNR, and total force (ALNAV) data, the latter being the sum of the other two files. Thus, snapshots of those forces over time exist to which time series analysis can be applied. Also used by NAPPE for validation and forecasting purposes are (1) the pay tables (e.g., Figure 2) that have been in effect since 1957, (2) managerial inputs in the form of pay grade totals for each quarter that is to be predicted, and (3) actual (historical) pay grade totals. If predictions are to be reliable, all of the data base must be kept as up-to-date as possible.

Objective

The objective of this report was to provide a detailed description of the analysis and procedures used to formulate the NAPPE model.

APPROACH

The basic approach taken by NAPPE is an aggregative one. Although the use of the historical quarterly inventories allows ready observation of personnel flows, making direct predictions for each pay grade by LOS cell would require additional knowledge relating to such policy variables as promotions, demotions, and occupational skills. This information, in turn, would not only lead to the expansion of the data base but also to a more deterministic model. Instead, NAPPE relies solely upon historical data to project the future. It forecasts the LOS marginal distribution by considering its statistical properties and then distributes the personnel over the authorized pay grade totals to arrive at the final force structure matrix. The following sections describe in detail the methods employed to forecast future inventories.

Forecasting the LOS Distribution

Examination of the quarterly force structure tables suggests two alternative approaches for predicting LOS cells 2 through 31. (Cell 1 is a special case and will be discussed later.) The first involves the use of the actual cell populations, denoted $P_j(t)$ (i.e., the population with longevity j at time t, with t being measured in quarters). As will be demonstrated shortly, a forecast for LOS cell j at time t, $F_j(t)$, is calculated utilizing actual and/or previous forecasts for the quarters through time t-1. That is,

$$F_j(t) = \hat{P}_{j(t-1)}. \qquad (1)$$

The second approach consists of using a set of transformed variables that measure the transitivity of the force. They are defined as net loss rates (r) and are given by

$$r_j(t) = (P_j(t) - P_{j+1}(t+m))/P_j(t), \qquad (2)$$

for j = 2,3,...,30. (The loss rate for cell 31 is also a special case and is discussed in a later section.) Thus, $(1-r_j(t))$, or $C_j(t)$, is the continuance or transition rate; namely, the proportion of personnel in LOS cell j at time t that "move" to cell j + 1 at time t + m. Since length-of-service is measured in years, m is set equal to 4 in NAPPE, thus providing yearly transition rates. A forecast for LOS cell j + 1 at time t is then computed as

$$F_{j+1}(t) = \hat{C}_j(t-4)P_j(t-4) \qquad (3)$$

where $\hat{C}_j(t-4)$ is the predicted rate of flow from $P_j(t-4)$ to $P_{j+1}(t)$. Because $P_j(t-4)$ is known, this method requires only the forecasting of the continuance rates.

Given the two data sources available in the USN, USNR, and ALNAV files, there are four straightforward ways of forecasting the ALNAV population of LOS cell j + 1 at time t using the two approaches outlined above. They can be written as

$$F^1_{j+1}(t) = P_{j1}(t-4)\hat{C}_{j1}(t-4) + P_{j2}(t-4)\hat{C}_{j2}(t-4),$$

$$F^2_{j+1}(t) = P_{j3}(t-4)\hat{C}_{j3}(t-4),$$

$$F^3_{j+1}(t) = \hat{P}_{j+1,1}(t-1) + \hat{P}_{j+1,2}(t-1), \qquad (4)$$

or

$$F^4_{j+1}(t) = \hat{P}_{j+1,3}(t-1),$$

where $P_{jk}(t)$ designates the actual number of personnel in cell $\underline{j}$ at time $\underline{t}$ for data set k = 1,2,3, indicating USN, USNR, and ALNAV, respectively. Note that the first two forecasts use predicted continuance rates, whereas the latter two use actual population projections. Also, the first and third forecasts combine USN and USNR predictions to give an ALNAV forecast.

Time series analysis techniques (cf. Box & Jenkins, 1970; Brown, 1963) were applied to the three data sets to forecast the continuance rates. For each of the 90 time series (30 different loss rates for each of the three data files), each consisting of 72 data points, the rates were plotted, along with their autocorrelation and partial autocorrelation functions. Also, the first differences of the loss rates and their autocorrelation and partial autocorrelation functions were computed and plotted. Seasonality factors were identified by calculating the autocorrelation functions and their standard errors. Finally, a chi-square test was made on the first 24 autocorrelations to determine whether they could be distinguished from white noise. This procedure was used in an attempt to identify a general class of time series models that would be applicable to most or all of those of the individual time series while, at the same time, being conservative in the consumption of computer time.

Table 1 presents the series' seasonality factors and suggests an appropriate time series model based on the autocorrelation and partial autocorrelation functions. The terminology employed is that of Box and Jenkins (1970, Chapters 3 and 4), where a (p,d,q) autoregressive integrated moving average (ARIMA) model is differenced $\underline{d}$ times and contains $\underline{p}$ autoregressive and $\underline{q}$ moving average parameters. Thus, an entry in the table such as "(1,1,0) x 4" implies that a (1,1,0) model be fitted to the series created by taking $Z_t - Z_{t-4}$ for $t \geq 5$, where Z_t is the t^{th} observation in the series. This step eliminates a seasonality factor of period 4.

Table 1

Seasonality Factors and Appropriate Models for Loss Rate Time Series

	USNR		USN		ALNAV	
LOS	Seasonality Factor	Appropriate Model	Seasonality Factor	Appropriate Model	Seasonality Factor	Appropriate Model
1	2[a]	(0,1,0) x 2	2	(0,1,0) x 2	4[a]	(0,1,0) x 4
2	NONE	(0,1,0)	NONE	(1,1,0);(0,2,0)	NONE	(0,1,0)
3	2	(0,1,0) x 2	2[a]	(0,1,0) x 2	2[a]	(0,1,0) x 2
4	NONE	(0,1,1)	2[a]	(0,1,0) x 2	2[a]	(0,1,0) x 2
5	NONE	(1,0,1)	2[a]	(0,1,0) x 2	2[a]	(0,1,0) x 2
6	NONE	(0,1,0)	NONE	(0,1,1);(1,1,0)	NONE	(0,1,1);(1,1,0)
7	NONE	(2,0,0)	NONE	(0,1,0)	NONE	(0,1,1)
8	NONE	(0,0,2)	NONE	(0,1,0)	NONE	(0,1,1)
9	NONE	(2,0,0)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
10	NONE	(1,0,0)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
11	NONE	(1,0,0);(2,0,0)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
12	NONE	(0,0,1)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
13	NONE	(0,0,1)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
14	NONE	(0,0,1)	4[a]	(0,1,0) x 4	NONE	(0,1,0)
15	NONE	(0,0,1)	NONE	(0,1,0)	NONE	(0,1,0)
16	NONE	(0,0,1)	NONE	(0,1,0)	NONE	(0,1,0)
17	NONE	(0,0,1)	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
18	NONE	(0,0,1)	NONE	(0,1,0)	NONE	(0,1,0)
19	NONE	(2,0,0)	NONE	(0,1,0)	NONE	(0,1,0)
20	3[a]	(0,1,0) x 3	NONE	(0,1,0)	NONE	(0,1,1)
21	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4
22	NONE	(1,0,0)	NONE	(0,1,0)	NONE	(0,1,0)
23	NONE	(1,0,0)	4[a]	(1,1,0) x 4	NONE	(2,0,0)
24	NONE	(1,0,0)	NONE	(0,1,0)	NONE	(0,1,0)
25	NONE	(0,1,0)	NONE	(0,1,0)	NONE	(0,1,0)
26	NONE	(1,0,0)	3[a]	(0,1,0) x 3	3[a]	(0,1,0) x 3
27	NONE	(1,0,0)	NONE	(0,1,0)	NONE	(0,1,0)
28	NONE	(1,0,0)	2[a]	(0,1,0) x 2	NONE	(0,1,0)
29	NONE	(0,1,1)	2[a]	(0,1,0) x 2	2[a]	(0,1,0) x 2
30	4	(0,1,0) x 4	4[a]	(0,1,0) x 4	4[a]	(0,1,0) x 4

[a] The first 24 autocorrelations of (0,1,0) model were not significantly different from white noise, although one (or more) of them is significant by itself.

The undifferenced data in all of the USN and ALNAV series and in some of the USNR series indicated nonstationary behavior. A large number of these appear to be well suited to a single exponential smoothing model because, in general, they do not vary about a fixed mean, yet they exhibit homogeneous behavior of a kind. In this regard, Box and Jenkins (1970) comment: ". . . although the general level about which fluctuations are occurring may be different at different times, the broad behavior of the series, when differences in level are allowed for, may be similar" (p. 11). The exponential smoothing model is most easily recognized by its recursion formula,

$$\overline{Z}_t(\alpha) = \alpha Z_{t-1} + (1-\alpha)\overline{Z}_{t-1}(\alpha) \quad , \text{ for } 0 \leq \alpha \leq 1 \tag{5}$$

which shows that each new level of the process (the forecast), $\overline{Z}_t(\alpha)$, is arrived at by interpolating between the new observation, Z_{t-1}, and the previous level, $\overline{Z}_{t-1}(\alpha)$. With α at its upper limit of 1, the level of the process is simply the latest observation, with all previous history ignored. With α small (closer to 0), the previous level is given more weight than the latest observation. The smaller α is, the longer the time period required by the model to recognize a change in "levels" of the data.

Thirty-two of the time series seem to be modelled well by single exponential smoothing. First, 25 of the series were reduced to white noise after first differencing, implying a (0,1,0) model (i.e., $Z_t = Z_{t-1} + a_t$ where a_t is the noise component and Z_t is the t^{th} observation of the series). But this is simply equation (5) with α equal to 1. As shown in Table 1, these series were USNR cells 2, 6, 25; USN cells 7, 8, 15, 16, 18-20, 22, 24, 25, 27; and ALNAV cells 2, 14, 15, 16, 18, 19, 22, 24, 25, 27, 28. It was indicated that another seven series (USNR cells 4, 29; USN cell 6; and ALNAV cells 6, 7, 8, 20) could be fitted by a (0,1,1) ARIMA process (i.e., $Z_t = Z_{t-1} + a_t - \Theta a_{t-1}$, where $\Theta = 1-\alpha$). However, since this is the inverted form of equation (5) (see Box & Jenkins, pp. 105-106), these series can also be forecasted using the single exponential smoothing model. Note that double and triple exponential smoothing are not applicable here because no evidence exists of any linear or quadratic trends in the data.

Another 36 series, which required first differencing to achieve stationarity, contained evidence of a seasonality factor. A significant autocorrelation at lag 2 was noticed in 12 of these series (USNR cells 1, 3; USN cells 1, 3, 4, 5, 28, 29; and ALNAV cells 3, 4, 5, 29). Here, a significant autocorrelation at lag n means that pairs of observations separated by n time periods are significantly correlated. Twenty-one series possessed an indication of a cycle of length 4 (USNR 21, 30; USN 9-14, 17, 21, 23, 30; and ALNAV 1, 9-13, 17, 21, 30).

Finally, a significant autocorrelation at lag 3 was noted for three series (USN and ALNAV cell 26, USNR cell 20). The autocorrelations at lag 2 and lag 4 are most likely explained by the Navy's recruiting policies, which contain

an annual (or semiannual) pattern, but no explanation can be made for the autocorrelations at lag 3. Although the first differences of 33 of these 36 series ("a" in Table 1) cannot be distinguished from white noise (and hence a (0,1,0) model possibly would be appropriate), it was felt that all of the information in these series was not being utilized. Further research, including model fitting and diagnostic checking, verified this hypothesis. Thus, two seasonal models based on first differences, one accounting for period 2 and one for period 4, were included in the set of models to be used for the loss rate time series. The model for period 2 is written $Z_t = Z_{t-1} + Z_{t-2} - Z_{t-3} + a_t$; and that for period 4, $Z_t = Z_{t-1} + Z_{t-4} - Z_{t-5} + a_t$.

Because the single exponential smoothing model and the two seasonal models adequately explain 9 of the USNR series, 28 of the USN series, and 28 of the ALNAV series, no further models for the loss rates were instituted in NAPPE. The remaining 25 series (21 of them being USNR) not covered by these models are explained by an assortment of various models, including the (0,0,1), (1,0,0), (2,0,0), (1,0,1), (0,0,2), and (1,1,0) ARIMA processes. Since each of these models requires the estimation of at least one parameter, increased forecasting accuracy was sacrificed in the interest of using less computer time. It should be noted that the α parameter in the exponential smoothing model is estimated by choosing that α from the set $\{0, .05, .10, \ldots, .95, 1.0\}$ that minimizes historical error. In effect, then, 21 exponential smoothing models (for 21 α's) and 2 seasonal models (requiring no parameter estimation) are fitted to <u>all</u> 90 time series and the model that best explains the past is chosen to forecast future transition rates.

Similar procedures were used to produce a set of models that would fit most of the actual population data. Table 2 summarizes the analysis for the 90 time series covering LOS cells 2-31. (Cell 1 is a special case and will be discussed later.) Of special interest is the indication that 55 of the 60 USN and ALNAV series possess a seasonality factor of 4; and that all but 2 of these series (USN and ALNAV cell 31) are reduced to white noise by a (0,2,0) x 4 model. This model is written $Z_t = 2Z_{t-1} - Z_{t-2} + Z_{t-4} - 2Z_{t-5} + Z_{t-6}$, so that the six latest data points are required to make a new forecast. Another model that was chosen for the actual population data is the (0,1,0) x 4, which has previously been described. Also, the single exponential smoothing model is included because 9 of the USNR series could be fitted well by either a (0,1,0) or (0,1,1) model.

No other models were selected, primarily because most of the series were covered by the models already chosen. Also, other models that might have been included, such as a (2,0,0) ARIMA process, require far too much parameter estimation (with values ranging from -1 to 1) to make them worth the effort. Finally, it was hypothesized that NAPPE would consistently be using the forecasts provided by the continuance rates time series because of their relative stability as compared to the actual population series. The latter, as expected, contained much more variability because of the rise and decline in force size during the 1960s. The hypothesis was later verified during the execution of the model.

Table 2

Seasonality Factors and Appropriate Models for Actual Population Time Series

LOS	USNR Seasonality Factor	USNR Appropriate Model	USN Seasonality Factor	USN Appropriate Model	ALNAV Seasonality Factor	ALNAV Appropriate Model
2	NONE	(0,1,1)	4[a]	(0,2,0) x 4	4[a]	(0,2,0) x 4
3	4	(0,1,0) x 4	4	(0,2,0) x 4	4	(0,2,0) x 4
4	NONE	(0,1,1);(0,1,0)	4[a]	(0,2,0) x 4	4[a]	(0,2,0) x 4
5	NONE	(0,1,1)	NONE	(0,1,0)	NONE	(0,1,0)
6	NONE	(1,1,0)	NONE	(0,2,0)	NONE	(0,2,0)
7	NONE	(2,0,0)	4[a]	(0,2,0) x 4	4[a]	(0,2,0) x 4
8	NONE	(1,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
9	3[a]	(0,1,0) x 3	4	(0,2,0) x 4	4	(0,2,0) x 4
10	3[a]	(0,1,0) x 3	4	(0,2,0) x 4	4	(0,2,0) x 4
11	2	(0,1,0) x 2	4	(0,2,0) x 4	2	(0,2,0) x 2
12	NONE	(0,1,1)	4	(0,2,0) x 4	4	(0,2,0) x 4
13	NONE	(1,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
14	NONE	(1,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
15	NONE	(1,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
16	NONE	(1,1,0);(2,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
17	NONE	(1,1,0);(2,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
18	NONE	(1,1,0);(2,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
19	NONE	(1,1,0);(2,0,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
20	NONE	(1,1,0);(2,0,0)	4[a]	(0,2,0) x 4	4[a]	(0,2,0) x 4
21	NONE	(2,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
22	NONE	(1,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
23	NONE	(2,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
24	NONE	(2,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
25	NONE	(1,1,0)	NONE	(1,1,0)	NONE	(1,1,0);(2,0,0)
26	NONE	(0,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
27	NONE	(0,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
28	NONE	(0,1,0)	4	(1,1,0)x4;(0,2,0)x4	4	(0,2,0) x 4
29	4[a]	(0,1,0) x 4	4	(0,2,0) x 4	4	(0,2,0) x 4
30	NONE	(0,1,0)	4	(0,2,0) x 4	4	(0,2,0) x 4
31	NONE	(0,1,0)	4	(0,1,0) x 4	4	(0,1,0) x 4

[a]The first 24 autocorrelations of nonseasonal model were not significantly different from white noise.

USN, USNR, and ALNAV forecasts for a particular quarter are provided for each LOS cell using continuance rates and actual population data. These individual forecasts are produced by the model that minimizes historical error. These six forecasts for each cell (two forecasts for a particular LOS cell in each of three data sets) are used to produce the four ALNAV forecasts seen in equation (4), where the USN and USNR forecasts are combined.

Examination of the forecasts F^1 through F^4 in equation (4) showed that, first, when an influx of USNR personnel with prior service occurs, some ALNAV loss rates may be negative. In those cases, F^1, which is based on the sum of the individual predictions of USN and USNR, is more accurate than the direct ALNAV forecast, F^2. Also, as was mentioned earlier, the forecasts based on continuance rates (F^1, F^2) yielded less historical error than those based on actual population data (F^3, F^4). Consequently, in some cases, the latter forecasts could be used to "damp" the continuance rate forecasts. Such instances would be characterized by conditions in which the continuance rate varies inversely with the source population. Thus, a fifth prediction was made by taking a weighted sum of the first four:

$$F^5_{j+1}(t) = \beta[F^1_{j+1}(t) + F^2_{j+1}(t)]/2 + (1-\beta)\ [F^3_{j+1}(t) + F^4_{j+1}(t)]/2, \tag{6}$$

$0 \leq \beta \leq 1$. Here β is some weight that produces a minimum sum of squared error. Finally, the forecast for a particular cell is the one chosen from $\{F^1, F^2, F^3, F^4, F^5\}$ which has been the best estimator of the past, with "best" implying the forecast that minimizes historical error. As will be seen, this "final" forecast will most likely be modified.

Special Considerations for Cells 31 and 1

As noted earlier, LOS cell 31 contains all personnel with more than 30 years of service. Because the continuance rate as previously defined only explains the movement of personnel from cell 30 to cell 31, a special rate was defined that would encompass the movement of personnel from cell 30 as well as the continuance of those personnel already in cell 31. It is written

$$r_{31}(t) = [(P_{30}(t) + P_{31}(t)) - P_{31}(t+4)]/(P_{30}(t) = P_{31}(t)). \tag{7}$$

(Compare with (2).) The time series composed of these rates was analyzed in the same fashion of those for $r_2(t)$ through $r_{30}(t)$.

Cell 1 is also a special case; that is, no continuance rate can be defined for it because of the lack of a source population with military service less than 0 years. Also, due to the fluctuating input of recruits (determined largely by personnel policies), the high degree of instability of populations in cell 1 make it undesirable to smooth the historical populations in that cell. Yet, one ready-made forecast is available from the realization that the sum of the forecasts for cells 1 through 31 must equal the total force size.

Thus, we have the "residual forecast:"

$$\hat{P}_1(t) = F(t) - \sum_{j=2}^{31} \hat{P}_j(t), \tag{8}$$

where $\hat{P}_1$ is the ALNAV forecast for cell 1, F is the total force, and $\hat{P}_j$ is the best forecast selected from $\{F^1, F^2, F^3, F^4, F^5\}$ for LOS cell j. However, this forecast for cell 1 may be abnormally small, large, or even negative due to large changes in the total force size and the relative stability of forecasts $\hat{P}_2$ through $\hat{P}_{31}$. Therefore, the following analysis was performed to obtain more realistic forecasts for all the $\hat{P}_j$'s.

First, proportions were computed that relate the size of cell one to the total force size over the historical data base. Symbolically, these ratios, denoted R_i, are determined by

$$R_i = P_1(i)/F(i) \quad , \text{ for } i = 1, \ldots, 76 \tag{9}$$

where _i_ represents 76 quarters of history and $P_1(i)$ and F(i) are the cell 1 and total force populations for quarter _i_. Because the primary matter of concern is forecasting, the percent changes of these ratios over time, R_{i+1}/R_i, were examined. This series of 75 percentage changes indicates how the size of cell 1 relative to the total force changes over time. The mean of this series is 1.007, with a standard deviation of .08. Based on this information and on the assumption that the percentage changes possess a normal distribution (there is no evidence to the contrary), the following statement can be made:

$$\Pr\{0.85{*}R_i < R_{i+1} < 1.16{*}R_i\} \simeq .95. \tag{10}$$

In other words, the probability is approximately 95 percent that the ratio of cell 1 to the total force in quarter i + 1 is between 0.85 and 1.16 times the ratio in quarter _i_. Because the total force size for any projected quarter is known (based on authorized end strength), this method allows upper and lower bounds to be placed on the size of cell 1.

Another estimate for the actual size of cell 1 may be seen in the fact that the percentage changes R_{i+1}/R_i possess a correlation of .82 at lag 1. This implies that, even though the ratio of cell 1 to the total force may fluctuate over time, it is fairly stable over a short period. Thus, the actual population for cell 1 may be estimated by applying the previous quarter's ratio to the known total force size for the quarter being forecasted. In NAPPE, only the last historical quarter's ratio is used in forecasting all future quarters; significant error would be introduced if forecasted ratios were used instead. This method has worked out well, although difficulties may arise if the last quarter's ratio happens to be a turning point (up or down) in the series of cell 1 ratios.

The final formula for computing the cell 1 population for any future quarter t uses the residual forecast $\hat{P}_1$ for that quarter as well as the above procedures:

$$\hat{P}^*_1(t) = \max \{0.85*R_i*F(t), \min \left(\frac{\hat{P}_1(t) + (R_i*F(t))}{2}, 1.16*R_i*F(t) \right) \} \tag{11}$$

where the first term in the brackets is the lower bound on the size of cell 1; the second, the average of the residual forecast and the "last ratio" forecast; and the third, the upper bound on the size of cell 1. Here the subscript i refers to the last known quarter of historical data.

However, it is quite likely that the adjusted cell 1 forecast, $\hat{P}^*_1(t)$, is different from $\hat{P}_1(t)$, the residual forecast. Thus, the discrepancy, $\hat{P}_1(t) - \hat{P}^*_1(t)$, must be distributed over each of the remaining cells. The procedure used ensures that the cells that have been predicted with a high degree of accuracy are altered only slightly. It distributes the discrepancy according to the amount of the prediction error in a cell relative to the total prediction error over all other cells:

$$\hat{P}^*_j(t) = \hat{P}_j(t) + (\hat{P}_1(t) - \hat{P}^*_1(t))*(E_j(t)\hat{P}_j(t)/\sum_{j=2}^{31} E_j(t)\hat{P}_j(t)), \tag{12}$$

$j = 2,3,\ldots,31$, where the error is calculated as

$$E_j(t) = \sum_{i=5}^{t-1} |P_{j3}(i) - \hat{P}_j(i)|/P_{j3}(i) \tag{13}$$

for cell j. Recall that $P_{j3}(i)$ denotes the actual ALNAV population for cell j in quarter i. Also note that four quarters are required to initiate the forecasting process (for computation of continuance rates); hence, the lower range in the summation.

Having executed the above procedure, NAPPE now has the LOS marginal distribution $(\hat{P}^*_1, \hat{P}^*_2, \hat{P}^*_3, \ldots, \hat{P}^*_{31})$ of the force structure matrix. The vector is complete and sums to the total authorized strength. The pay grade vector is either supplied by the user or interpolated by NAPPE (discussed in a later section). Thus, the problem now is to calculate the interior entries in the force structure matrix. Once this has been accomplished, multiquarter forecasts are obtained by concatenating the predicted quarter with the historical data base and employing the optimal time series models that were chosen for the first prediction.

Constructing the Force Structure Matrix

Given the two marginal distributions of the force structure matrix, NAPPE follows Mosteller's (1968) procedure for deriving the interior. This interative procedure involves the renormalization of the rows and columns of a standard matrix to obtain the given marginals, while maintaining the relative

"associations" among the entries of the standard matrix. Although the final distribution is affected more by the configuration of the standard matrix than by its particular entries, analysis performed by Wicker (Note 2) showed that some improvement could be gained by using a weighted average (vice a straightline average) of historical force structures. Thus, NAPPE calculates the standard matrix S^1_{ij}, $i = 1,2,\ldots,31$, $j = 1,2,\ldots 9$, from

$$S^1_{ij} = .31192A^1_{ij} + .04583A^2_{ij} + .64225A^3_{ij}, \tag{14}$$

where A^1_{ij} is the present yearly average inventory matrix, A^2_{ij} is the yearly average 1 year in the past, and A^3_{ij} is the yearly average 2 years in the past.

Suppose now that the LOS and pay grade marginals are denoted by L_i and G_j, respectively, where $i = 1,2,\ldots,31$, and $j = 1,2,\ldots,9$. With these and the standard matrix (14) being given, Mosteller's procedure involves the following steps:

1. Let $n = 1$.
2. Let $S^{n+1}_{ij} = (L_i/\sum_j S^n_{ij})S^n_{ij}$ for all $\underline{i}$.
3. Let $S^{n+2}_{ij} = (G_j/\sum_i S^{n+1}_{ij})S^{n+1}_{ij}$ for all $\underline{j}$.
4. If $\sum_j S^{n+2}_{ij} \neq L_i$ for all $\underline{i}$ go to (2), otherwise stop.

This process allows estimates of a force structure matrix to be made for specific points in time. To predict basic pay costs for a fiscal year, some estimate of the configuration of that matrix throughout the year is necessary.

<u>Estimating Average Strength</u>

Given an actual inventory at some quarterly interval and a forecast inventory representing a period four quarters later, an estimate of average strength can be computed; that is,

$$\overline{S}_{ij} = (S_{ij} + \hat{S}_{ij})/2, \tag{15}$$

where S_{ij} and $\hat{S}_{ij}$ represent the actual and forecasted strength matrix entries, respectively. Clearly, the average force $\overline{S}_{ij}$ derived in this manner assumes that changes in a force configuration over four quarters occur in a constant relation to time. Thus, if the difference between $\sum\sum_{ij} S_{ij}$ and $\sum\sum_{ij} \hat{S}_{ij}$ were 100, an average matrix would reflect a monthly change of 8.33 or a quarterly change of 25, etc.

However, straightline methods are inadequate because changes do <u>not</u> occur uniformly throughout the year. Unfortunately, the only practical alternatives in terms of available data are average strength computations based on quarterly increments. Thus, given S_{ij} for t = December 1975, a forecast of $\hat{S}_{ij}$ is made for t = December 1976. Similarly, four-quarter forecasts are made June to June, September to September, and March to March. So, by forming force structure matrices for each quarterly prediction and by combining those four forecasts with the most recent actual matrix, a relatively good estimate of a yearly average strength can be computed by a five-point average:

$$\bar{S}_{ij} = (S_{ijt_0} + \hat{S}_{ijt_1} + \hat{S}_{ijt_2} + \hat{S}_{ijt_3} + \hat{S}_{ijt_4})/5. \tag{16}$$

Up to this point, estimates of future pay grade distributions were assumed to be an exogenous input accessible to the users of the model. However, the estimation of the nine cells of the pay grade marginal distribution is no trivial task.

While the end-year pay grade structure is roughly established by the budget (authorized end strength), the quarterly pay grade distributions are largely a function of management actions taken throughout the fiscal year. If pay grade totals for September, December, March, and June are input by the user, the model will form complete force structure matrices based on those given pay grade marginals and the forecasted longevity marginals. This input will be based on official plans that reflect management intentions. Specifically, the pay grade populations for the six petty officer grades (E-4 through E-9) mainly reflect the planned phasing of petty officer promotions during the subject year, while the three lowest grades reflect the planned phasing of recruit and reserve input as modified by attrition and differential rates of promotion from E-1 to E-2 and E-2 to E-3.

If, instead, the pay grade totals for the three quarters interior to the fiscal year (the beginning total is known and the ending total is always input as a matter of strength/budget policy) are to be forecasted by NAPPE, the following procedures are used: Let $\hat{F}_{t,i}$ and $\hat{F}_{t+4,i}$ be the actual or user input pay grade totals for the end (or beginning) of 2 consecutive fiscal years. Estimates of the quarterly pay grade totals are then given by

$$\hat{F}_{t+2,i} = (\hat{F}_{t,i} + \hat{F}_{t+4,i})/2,$$

$$\hat{F}_{t+1,i} = (\hat{F}_{t,i} + \hat{F}_{t+2,i})/2, \tag{17}$$

and

$$\hat{F}_{t+3,i} = (\hat{F}_{t+2,i} + \hat{F}_{t+4,i})/2.$$

These simple linear interpolations could be altered by increasing or decreasing them from the straightline results by reference to an average of historical seasonal patterns of pay grade population movement. However, analysis revealed wide variability in the quarterly deviations from straightline averages from year to year. Thus, the decision was made to allow the user to implement any seasonal pattern he feels is appropriate by directly inputting the quarterly future pay grade totals. Otherwise, NAPPE will use those obtained in equation (17).

Costing the Force Structure

The procedures outlined previously produce a forecast of the force structure matrix for the end of each quarter (30 September, 31 December, 31 March, and 30 June) during the forecast year. In addition, the beginning matrix (30 June) is available in the form of known (i.e., historical) data.

The following costing procedures are observed. If the pay scale is constant for the year, the five matrices or inventories can be averaged and a pay scale such as that in Figure 2 applied to the average force to produce a forecast of annual basic pay costs. If the pay scale changes during a fiscal year, some provision must be made for costing part of the average strength with one pay scale and for costing its complement with another. Since no data are available to forecast population changes within a quarter, it is assumed that intraquarter changes occur uniformly throughout the quarter. Also, if the pay scale is in effect for a full quarter, the average strength for that quarter is defined by the average of the begin and end quarter matrices, and the cost is derived by applying the constant pay rates to the quarterly average. Finally, if the pay scale changes during the quarter, then (1) the force structure at the point of change is approximated by linear interpolation between the begin and end quarter matrices, (2) the populations for the start and end of each pay scale are averaged to obtain the average force to which the relevant pay scale is applied, and (3) a corresponding daily pay scale is applied to the average force. The results then are multiplied by the number of days for which the pay scale is applicable. The results for all quarters--whether single or dual pay scales--are summed over the fiscal year to produce an estimate of annual basic pay.

This method is followed for any fiscal year to be predicted. It requires only the end strength pay grade totals for each year to be forecasted. The user may supply quarterly pay grade totals as desired.

RESULTS

Historical Validation of the NAPPE Model

One problem in obtaining a complete validation of NAPPE lies in the fact that the authorized end strengths are seldom achieved exactly. Consequently, although some discrepancy may exist between the actual and the forecasted force structures, it would not significantly affect the validation results.

To demonstrate NAPPE's accuracy, the model was used to forecast historical force structures and their respective costs for the fiscal years 1965-1976. NAPPE was allowed to use historical data up to 1, 2, and 3 years preceding the year to be predicted. All pay grade totals used were actual quarterly totals as opposed to authorized end strength totals, which may or may not have been achieved. It should be noted that, although the forecast for any given cell in the force structure matrix could be erroneous, the total cost may be accurate due to offsetting errors in other cells.

Therefore, an additional validation criterion was established to test for systematic bias in the forecasting process. This involved comparing the forecasted and actual longevity distributions, with the criterion being average length of service. Table 3 summarizes these results, given the three different lead times. As shown, there appears to be no evidence of consistent under- or overestimation of mean LOS, although, in general, the forecasts are less accurate as lead time increases.

The historical validation for basic pay involved the comparison of the costs of the projected force structure and the actual force structure. The absolute values of the percentage differences between the two also are included in Table 3, which demonstrates NAPPE's accuracy for forecasts made 1, 2, and 3 years in advance of the projected period. Obviously, the Viet Nam years were most difficult to forecast as the size and composition of the enlisted force fluctuated greatly. However, the pay predictions for the later years are very close to the actual costs, with the last eight quarters in the table being predicted within .2 percent for all three lead times. Overall, the average 1-year lead time percent error in cost prediction in Table 3 is .24 percent, with a standard deviation of .26 percent. Similar figures for the 2-year and 3-year lead time forecasts are .76 (.78) and 1.30 (1.06), respectively. Once again, accuracy generally diminishes with an increase in lead time.

The Implementation of NAPPE

NAPPE is currently operating in a demand mode on a UNIVAC 1110 computer located at the Naval Ocean Systems Center, San Diego and in a batch mode on an IBM 360 at the Bureau of Naval Personnel (BUPERS), Washington, D.C. In the later instance, NAPPE is used primarily in two ways.

Table 3

Validation: Predicted Mean LOS and Percent Error in Cost Prediction Based on Actual Quarterly Pay Grade Totals

Lead Time Quarter	1 Yr. Lead Time Predicted Mean LOS	1 Yr. Lead Time % Error Cost Pred.	2 Yrs. Lead Time Predicted Mean LOS	2 Yrs. Lead Time % Error Cost Pred.	3 Yrs. Lead Time Predicted Mean LOS	3 Yrs. Lead Time % Error Cost Pred.	Actual Mean LOS
9/64	6.05	.10					6.04
12/64	6.08	.18					6.07
3/65	6.06	.25					6.07
6/65	6.04	.29					6.05
FY 65	6.06	.21					6.06
9/65	5.98	.11	5.96	.16			5.91
12/65	5.89	.29	5.89	.01			5.76
3/66	5.90	.47	5.88	.10			5.72
6/66	5.85	.81	5.85	.39			5.63
FY 66	5.93	.43	5.92	.10			5.81
9/66	5.63	.13	5.80	1.04	5.80	.70	5.58
12/66	5.55	.19	5.84	1.33	5.84	1.02	5.56
3/67	5.52	.10	5.86	1.76	5.84	1.43	5.53
6/67	5.45	.03	5.86	2.00	5.86	1.66	5.51
FY 67	5.55	.10	5.84	1.53	5.84	1.21	5.56
9/67	5.48	.05	5.45	.19	5.81	1.74	5.51
12/67	5.49	.29	5.40	.47	5.89	1.49	5.47
3/68	5.43	.59	5.36	.63	5.89	1.72	5.39
6/68	5.39	.74	5.30	.69	5.85	1.83	5.35
FY 68	5.46	.43	5.39	.50	5.86	1.70	5.45

Table 3 (Continued)

Lead Time Quarter	1 Yr. Lead Time Predicted Mean LOS	1 Yr. Lead Time % Error Cost Pred.	2 Yrs. Lead Time Predicted Mean LOS	2 Yrs. Lead Time % Error Cost Pred.	3 Yrs. Lead Time Predicted Mean LOS	3 Yrs. Lead Time % Error Cost Pred.	Actual Mean LOS
9/68	5.37	.09	5.44	.87	5.32	.68	5.34
12/68	5.39	.20	5.47	1.04	5.31	.50	5.38
3/69	5.32	.18	5.41	1.23	5.29	.23	5.30
6/69	5.23	.00	5.38	1.38	5.24	.08	5.25
FY 69	5.33	.12	5.42	1.13	5.29	.33	5.32
9/69	5.26	.04	5.30	.04	5.49	1.64	5.26
12/69	5.39	.01	5.41	.39	5.54	1.75	5.38
3/70	5.38	.12	5.45	.49	5.54	1.52	5.45
6/70	5.38	.60	5.39	.05	5.63	1.29	5.58
FY 70	5.33	.19	5.36	.22	5.52	1.55	5.38
9/70	5.63	.02	5.42	.94	5.48	.39	5.63
12/70	5.71	.21	5.55	.94	5.58	.37	5.78
3/71	5.78	.41	5.55	1.05	5.63	.49	5.85
6/71	5.89	.43	5.55	1.76	5.56	1.20	6.00
FY 71	5.72	.27	5.49	1.18	5.53	.62	5.77
9/71	6.05	.19	6.01	.29	5.59	2.35	6.05
12/71	6.07	.26	5.95	.17	5.67	2.13	6.04
3/72	6.11	.15	5.96	.45	5.66	1.18	6.12
6/72	6.25	.16	5.99	.62	5.64	1.47	6.19
FY 72	6.10	.19	5.96	.40	5.62	1.74	6.08
9/72	6.10	.05	6.29	.07	6.13	.41	6.08
12/72	6.13	.10	6.31	.18	6.03	.31	6.12
3/73	6.30	.06	6.35	.03	6.06	.60	6.24
6/73	6.31	.06	6.45	.04	6.04	.85	6.29
FY 73	6.21	.07	6.33	.06	6.05	.55	6.18

Table 3 (Continued)

Lead Time Quarter	1 Yr. Lead Time Predicted Mean LOS	1 Yr. Lead Time % Error Cost Pred.	2 Yrs. Lead Time Predicted Mean LOS	2 Yrs. Lead Time % Error Cost Pred.	3 Yrs. Lead Time Predicted Mean LOS	3 Yrs. Lead Time % Error Cost Pred.	Actual Mean LOS
9/73	6.24	.08	6.20	.23	6.49	.08	6.26
12/73	6.29	.18	6.23	.30	6.50	.10	6.32
3/74	6.31	.20	6.31	.29	6.49	.04	6.34
6/74	6.36	.12	6.34	.15	6.56	.11	6.31
FY 74	6.30	.14	6.28	.26	6.50	.08	6.30
9/74	6.24	.01	6.27	.08	6.18	.24	6.23
12/74	6.20	.06	6.25	.02	6.15	.28	6.13
3/75	6.21	.07	6.25	.05	6.23	.18	6.21
6/75	6.22	.02	6.34	.09	6.32	.11	6.24
FY 75	6.23	.02	6.29	.06	6.24	.20	6.22
9/75	6.21	.01	6.20	.07	6.28	.06	6.22
12/75	6.21	.07	6.23	.03	6.34	.05	6.15
3/76	6.23	.18	6.16	.07	6.26	.07	6.17
6/76	6.24	.18	6.15	.03	6.31	.02	6.21
FY 76	6.22	.11	6.19	.00	6.30	.01	6.20
9/76 (TQ)	6.15	.03	6.11	.06	6.11	.05	6.14
12/76	6.05	.13	6.08	.02	6.08	.00	6.12
3/77	6.10	.19	6.14	.02	6.07	.03	6.16
6/77	6.15	.11	6.18	.08	6.07	.10	6.16

First, during the budget (Program Objectives Memoranda (POM)) cycle, the number of personnel in the 31 longevity categories is forecasted up to 7 years in the future, based on manpower requirements (authorized end strengths) specified by the user. Other cost models at BUPERS are then used to compute future basic pay. These results are turned over to the Navy Manpower Planning System, which revises its accession plans if the forecasted budget constraints are not met. This cycle is repeated until the manpower requirements are satisfied by the personnel allocation within the revised budget constraints.

Second, NAPPE's computation of the mean length-of-service allows BUPERS to monitor the age of the force. Thus, Navy management can be kept aware of upward or downward trends in experience of the force. Also, NAPPE is used to check other personnel models, such as the Force Structure Simulation Model (known as the FAST model) (Boller, Note 3), that do not possess quarterly forecast capabilities.

For general use, NAPPE subroutines allow the user to forecast LOS transition rates, to update the historical data base, and to validate the model with respect to data ranging back to 1957. Still other routines permit the examination of historical transition rates or the force structure matrices themselves.

CONCLUSIONS

1. Sufficient data are available to allow time series analysis techniques to be applied toward the development of a model to forecast the enlisted force structure and, thus, basic pay obligations.

2. A particular set of time series models exists that adequately explains a great majority of the data. This set of models provides a solid statistical basis for making force structure forecasts.

3. The model can be utilized by Navy management to provide future basic pay and mean LOS estimates.

RECOMMENDATIONS

Additional work to be performed on the NAPPE model itself should concentrate more on the computer program than on the theoretical foundation; many of the techniques already used in the model were found to be quite accurate and reliable. Hence, any additional improvements in NAPPE's accuracy are likely to come in very small increments, at high costs, and with great effort. For the moment, then, attention should be given to making NAPPE more responsive to user needs. This effort should include the designing of an interactive NAPPE so that Navy personnel managers will be able to quickly investigate alternative force configurations and their respective costs.

Second, NAPPE should be expanded to include forecasts of other enlisted budgetary components. For example, in forecasting the obligations required for the basic allowance for quarters (BAQ) category, a force structure matrix indicating the dependency status of personnel may be generated and tied into the various BAQ pay rates. The same general principle could be followed for other categories, such as family separation allowances, subsistence in kind, clothing allowances, proficiency pay, separation payments, etc. In general, any cost variable that can be related to average strength in the form of a pay grade by LOS matrix (or some subset thereof) may be forecast using the techniques described in this report.

Finally, the Navy should develop a Naval Pay Predictor Officer (NAPPO) Model to forecast officer basic pay. (Preliminary efforts toward this are now in progress.) Hopefully, the characteristics of the officer force structures will be akin to those of the enlisted force, differing perhaps only in degree. In any event, once NAPPO has been developed and proven reliable, it can be extended to cover other officer costs besides basic pay.

REFERENCES

Box, G., P., & Jenkins, G. M. Time series analysis forecasting and control. San Francisco: Holden Day, Inc., 1970.

Brown, R. G. Smoothing, forecasting, and prediction of discrete time series. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1963.

Mosteller, F. Association and estimation in contingency tables. Journal of the American Statistical Association, 1968, 63(321), 1-28.

REFERENCE NOTES

1. Chipman, M. D., & Wilson, L. R. A descriptive analysis of Budget Activity Two of the Military Personnel, Navy Appropriation (NPRDC Special Rep. 77-10). San Diego: Navy Personnel Research and Development Center, June 1977.

2. Wicker, F. Unpublished Memorandum entitled "Error Minimization in NAPPE," 18 November 1976.

3. Boller, R., Offir, J., & Silverman, J. A force structure simulation model (Working Paper). San Diego: Navy Personnel Research and Development Center, 1976.

DISTRIBUTION LIST

Assistant Secretary of the Navy (Manpower, Reserve Affairs and Logistics)
Principal Deputy Assistant Secretary of the Navy (Manpower and Reserve Affairs)
Chief of Naval Operations (OP-987H), (OP-964D), (OP-01CR)
Chief of Naval Personnel (Pers-10c), (Pers-04), (Pers-2B)
Chief of Naval Material (NMAT 08T244)
Chief of Naval Research (Code 450) (4)
Chief of Information (OI-2252)
Director of Navy Laboratories
Commandant of the Marine Corps (Code MPI-20)
Chief of Naval Education and Training (005), (N-5)
Commander, Naval Data Automation Command
Commanding Officer, Naval Aerospace Medical Institute (Library Code 12) (2)
Director Training Analysis and Evaluation Group (TAEG)
Office of Civilian Personnel
Personnel Research Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base
Occupational and Manpower Research Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base
Technical Library, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base
Flying Training Division, Air Force Human Resources Laboratory, Williams Air Force Base
Technical Training Division, Air Force Human Resources Laboratory, Lowry Air Force Base
Advanced Systems Division, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base
Program Manager, Life Sciences Directorate, Air Force Office of Scientific Research (AFSC)
Army Research Institute for the Behavioral and Social Sciences
Military Assistant for Training and Personnel Technology, ODDR&E, OAD(E&LS)
Director for Acquisition Planning, OASD(MRA&L)
Coast Guard Headquarters (G-P-1/62)
Defense Documentation Center (12)